Towards Formal Analysis of Insider Threats for Auctions

Florian Kammüller
Computer Science
Middlesex University London
United Kingdom
f.kammueller@mdx.ac.uk

Manfred Kerber
Computer Science
University of Birmingham
United Kingdom
M.Kerber@cs.bham.ac.uk

Christian W. Probst
Technical University of Denmark
Copenhagen, Denmark
cwpr@dtu.dk

ABSTRACT

This paper brings together the world of insider threats and auctions. For online-auction systems, like eBay, but also for high-value one-off auction algorithms as they are used for selling radio wave frequencies, the use of rigorous machine supported modelling and verification techniques is meaningful to prove correctness and scrutinize vulnerability to security and privacy attacks. Surveying the threats in auctions and insider collusions, we present an approach to model and analyze auction protocols for insider threats using the interactive theorem prover Isabelle. As a case study, we use the cocaine auction protocol that represents a nice combination of cryptographic techniques, protocols, and privacy goals suitable for highlighting insider threats for auctions.

Keywords

Insider Threat; Auctions; Formal Methods

1. INTRODUCTION

Formal analysis and verification techniques applying interactive theorem proving have been successfully used for the analysis of auctions [3], as well as for security protocols [23]. Moreover, these techniques have also been used for insider threat analysis [12].

In this paper, we approach insider threats on auctions formally building on the experiences from those earlier approaches in order to arrive at a meaningful framework for a rigorous mathematical analysis of auction insider threats using automated reasoning.

As a running example, we use a fictitious cocaine auction protocol introduced by Stajano and Anderson [25] as a simplified example to first scrutinize privacy issues in the, then upcoming, eBay auction system and to argue the case for physical broadcast to realize anonymity. Following the inductive approach by Paulson to model and analyze security protocols [23], we formalize the protocol exhibiting insider threats: we first show that the formal definition in the inductive approach suffices to exclude the "sweetheart deal". In a sweetheart deal, the seller and one of his friends agree before the auction that the seller sells the good at the price determined in the auction. The participants in the auction assume that they have a fair chance of winning the good, but actually they have not, since the seller and his friend have already agreed on a deal, using the participants of the auction for determining the price. To arrive at formalizing a collusion between bidders ("ringing") we need to extend the model slightly and tap into some of the concepts used in the Isabelle insider framework [12].

The technical contributions of this paper are (a) a formalization of the cocaine protocol using Isabelle's inductive approach including the formalization and proof of the absence of the sweetheart deal and the impossibility of excluding collusion of insiders (b) the extension of the inductive approach to auctions by expressing arbitrary numbers of rounds, broadcast messages, an anonymity layer, and by merging with the Isabelle insider framework. The technical contributions lead to a deeper understanding of the relationship between auctions and security protocols with regard to insider threats paving the way for a more substantial Isabelle insider framework integrating relevant parts of the inductive approach to model and verify auction systems in the presence of insider threats.

We first review the auction literature with a special regard to security attacks involving collusions. Vice versa we summarize how the insider threat community deals with collusion of insiders (Section 2). Section 3 introduces our running example, the cocaine auction protocol, before we present its formalization in the Isabelle inductive approach (Section 4). We finally discuss to what extent the formal approach is useful to express possible insider threats to auctions, whether the threats exhibited on the case study are representative and complete, and summarize challenges for future research on this (Section 6).

2. AUCTION ATTACKS AND COLLUDING INSIDERS

In this section, we want first to give a very brief introduction to auctions and then discuss one of the biggest problem of auctions, collusion.

Auctions come in different forms, for instance, there are so-called first price and second price auctions. In first price auctions, the winner is the bidder with the highest bid and has to pay the value of his or her bid[1]. In second price auc-

[1]In case of two or more bidders with the same highest bid certain tie breaking rules apply.

ACM ISBN 978-1-4503-2138-9.
DOI: http://dx.doi.org/10.1145/2995959.2995963

tions, the winner is again the bidder with the highest bid, but has to pay the value of the second highest bid. The latter is in some ways more complicated than a first price auction.[2] So why is there interest in second price auctions? Since it makes matters much easier for the bidders. Vickrey's theorem states that in a second price auction a bidder cannot do better than bidding what the object is actually worth to her. If a bidder bid more than the object is worth to her she would make a loss on winning. If she bid less than the object is worth to her then potentially she does not realize a potential gain. If, however, she bids exactly how much she values the object all this cannot happen.

In addition to the distinctions between first price and second price auctions, auctions come in single round auctions, or auctions with several rounds (either ascending or descending). New auctions are still designed, e.g., recently for mobile phone licenses in different countries.

According to one of the leading auctions designers [14, p.152] "the two issues that really matter [in auction design] are attracting entry and preventing collusion." The first issue is that if not sufficiently many agents attend an auction then this is bad news for the seller. In the preface, Klemperer mentions the case of "a German auction of three blocks of spectrum [for which only three bidders had turned up and] which therefore sold only at a tiny reserve price." The second issue is more interesting for the purpose of this paper (albeit related to the first. If the bidders collude it only looks as if there were many bidders, but actually all the colluding bidders should count only as one). [14, p.152]: "Ascending auctions allow bidders to use the early rounds to signal each other how they might 'collusively' divide the spoils, and if necessary, use later rounds to punish any rivals who fail to cooperate.[...] By contrast, a (first-price) *sealed-bid auction* provides no opportunity for either signalling or punishment to support collusion."

While it is not possible to send signals, it is still possible to collude in such an auction as well, be it a first price or second price auction.

Krishna [15, p.152] describes collusion in second price auctions. Assume you have a number of agents and a bidding ring (or cartel) among those. The cartel would determine the one among them who values the object most and only the high bidder would submit a serious bid, all the others would either submit nothing or something that is so low that there is no danger that it will bring the price up. The bidders outside the cartel are not affected by this. 'A bidding ring generates profits for its members, of course, by suppressing competition.' In the extreme case all bidders are part of the cartel and the object will be sold at the reserve price at the expense of the seller, who without the existence of the cartel would achieve a higher price.

A separate issue does occur in a scenario in which the auctioneer cannot be trusted, for instance, since he is on the side of one of the participants and abuses his privileged position, e.g., to provide information to some participants but not to others.

Vice versa within the insider threat community, the collusion of insiders has been recognized as a main pattern of insider threats. The CMU-CERT Insider Threat Guide [4] names the *Ambitious Leader* pattern as one of the four main patterns of insider threats. This pattern describes an outsider – the ambitious leader – that works together with (at least) two insiders in separate infrastructures thereby realizing an attack that would not normally be possible for any of the involved insiders on their own. This pattern is a collusion of insiders. We used this pattern to show that the Isabelle insider framework [12] is capable of expressing all known insider threats.

[2]In case of combinatorial auctions when several goods are auctioned at the same time, the determination of the winners and the prices to pay may be computationally very complex.

3. COCAINE AUCTION PROTOCOL

In this section we will first summarize the cocaine protocol as described in [25], then look at potential formalizations, formalize the protocol, and discuss possible attacks.

3.1 Protocol

"Several extremely rich and ruthless men are gathered around a table. An auction is about to be held in which one of them will offer his next shipment of cocaine to the highest bidder. The seller describes the merchandise and proposes a starting price. The others then bid increasing amounts until there are no bids for 30 consecutive seconds. At that point, the seller declares the auction closed and arranges a secret appointment with the winner to deliver the goods." [25]. This is the short introduction to the cocaine auction protocol given in the original paper. This example serves as a model for eBay-like auctions where trust is an issue. In the eBay model, the auction house could drive up the sale price since it asks the bidders to reveal their maximum amount they are prepared to pay. The users simply have to trust that eBay will not exploit that knowledge to drive up the price (which would be profitable for the auction house because it takes a commission which is a percentage of the sales price). The eBay "peer review" system in which users give each other reliability rating has proven to be a quite successful method to guarantee trust between sellers and buyers. However, trusting the auction house remains a problem. The cocaine auction protocol has been designed as an "exaggerated case that makes the trust issue unambiguous" [25]. There are several assumptions imposed on the cocaine auction protocol in order to minimize trust.

- Nobody trusts anybody else more than is strictly necessary.
- The people that take part in the auction all know each other (otherwise one of them could be a police agent).
- No-one that makes a bid wants to be identified to the other bidders nor to the seller.
- Nobody apart from the seller and the buyer should know who won the auction; even the seller should only find out the identity of the buyer when committing to the sale, i.e., at the time of exchanging the goods at the secret appointment.
- None of the participants should have to trust any of the other participants; in particular there should not be an independent judge or policeman. The protocol must be self-enforcing.

3.2 Possible Implementations

For the context of this paper, we just assume an *anonymous broadcast:* a mechanism for broadcasting messages to

participants without revealing the identity of the sender. This represents an anonymity layer that can be implemented by cryptographic techniques, for example, using the dining cryptographers algorithm [5], or by using physical broadcast short-range radio networking facility, e.g., Piconet. In fact, the latter possibility is the main point of Stajano's and Ross' paper [25] to advocate the use of physical broadcast to implement the anonymity layer. For the context of this paper, we abstract from the concrete implementation of this anonymity layer. In the formal description of the protocol in Section 4, we will instead rely on the inductive approach to protocol verification and use address spoofing as a means for the senders to hide their identity from the receivers. There are two important details that we need to keep in mind when considering the practical implementation of the protocol.

1. The seller needs a mechanism to identify the winner.
   - A potential problem with this is that anyone can come later and claim to have said "yes" (i.e., made a bid) in the winning round.
   - A solution to this is that such a "yes" message (bid) contains a one-way function of a secret nonce.
   - The seller will ask the winner to exhibit the original nonce.
2. At finish of the auction, the seller prefers to give a secret appointment to the winner.
   - "See you on Tuesday at 06:30 in the car park of Heathrow terminal 5" (rather than exchanging suitcases of cocaine for cash under the noses of all the losing bidders).
   - On the other hand, the identity of the buyer should not be revealed to the seller until the latter commits to the sale (in order to protect the winner from not getting the bid for other biases, e.g., since he is from the "wrong family").

3.3 Protocol in Alice-Bob notation

Assuming an anonymity layer in a first approximation, the protocol can be described as follows.

- Identity of the seller is known to buyers.
- Buyers' messages are anonymous; seller's are not; all messages are broadcast.
- The protocol is a succession of rounds i of simple bidding.
  - The seller announces bid price b_i of round i.
  - Buyers have up to 30 seconds to say "yes".
  - As soon as a buyer says "yes", he is winner of the round, w_i.
  - A new round starts.
  - If 30 seconds elapse in round i with no bid, winner is w_{i-1}.

The implementation of this protocol is given informally [25] based on the use of the Diffie-Hellman key-establishment algorithm [6]. For convenience, we briefly summarize the main idea of the Diffie-Hellman key-establishment algorithm here. This algorithm uses a prime number modulus p and a generator $g \in \mathbb{Z}_p$ known to sender and receiver A and B. In addition, A keeps a secret number $a \in \mathbb{N}$ and B a secret number $b \in \mathbb{N}$. The algorithm works in two phases, establishing the shared secret $g^{ab} \texttt{mod}\ p$ between A and B without them ever exchanging their secrets a and b in clear. In the first phase, A calculates $g^a \texttt{mod}\ p$ and sends it to B; B calculates $g^b \texttt{mod}\ p$ and sends it to A. It is computationally infeasible for large prime numbers p to get a or b from these because of the Discrete-Logarithm-problem. In the second phase, A calculates $(g^b \texttt{mod}\ p)^a \texttt{mod}\ p$; B calculates $(g^a \texttt{mod}\ p)^b \texttt{mod}\ p$. Now, both have the shared secret because modular arithmetic gives

$$(g^b \texttt{mod}\ p)^a \texttt{mod}\ p = g^{ba} \texttt{mod}\ p = g^{ab} \texttt{mod}\ p = (g^a \texttt{mod}\ p)^b \texttt{mod}\ p.$$

The security of the algorithm depends on the high complexity of calculating $g^{ab} \texttt{mod}\ p$ given g, p and the sent messages $g^a \texttt{mod}\ p$ and $g^b \texttt{mod}\ p$. This problem is known as the Diffie-Hellman-problem and seems intuitively related to the computationally intractable Discrete-Logarithm-problem. The Diffie-Hellman-problem has been proved to be equivalent (for certain cases) to the Discrete-Logarithm-problem [20].

To formalize the protocol in the semi-formal "Alice-Bob" notation we first give the assumptions.

- Generator g and module p are public auction parameters.
- Anonymous "yes" message of winner w_i is g^{x_i}.
- Seller uses his (random) secret y to send the secret appointment to final winner w_i encrypted with Diffie-Hellman key $g^{x_i y}$.
- Possible variants for disambiguation and conciseness are possible.
  - Succession of bid prices b_i is pre-specified (conciseness).
  - At beginning of round i, seller broadcasts the "yes" message $g^{x_{i-1}}$ of winner of previous round to arbitrate races.
  - Bidders should include the b_i in their "yes" messages.

Some extensions to the standard point-to-point messaging that is common in Alice-Bob-notation are needed to express the anonymous communication. Stajano and Ross introduce a dedicated notation, which we adapt here for simplicity slightly.

- $\mathcal{D}$ is the set of auction principals including the seller S with secret y.
- $?A_i$ represents an anonymous principal in round $i = 1, \ldots, n-1$ with secret number x_i.
- Winning round $n-1$.

The cocaine auction protocol can then be specified using Diffie-Hellman and the notations $\{a, b\}$ for message concatenation and $\{m\}_K$ for encryption of message m with key K.

0. $S \to \mathcal{D}$: $g^y \bmod p$
i. $?A_i \to \mathcal{D}$: $\{g^{x_i} \bmod p,\ b_i\}$
n. $S \to \mathcal{D}$: $\{b_i, \mathit{MeetingAppointment}\}_K$, $K = g^{x_{n-1} y} \bmod p$

3.4 Insider/Collusion Attacks

Stajano and Anderson state "[t]here are limits what can be achieved on the protocol level. It is always possible, [...], to subvert an auction when an appropriate combination of participants colludes against others". This collusion attack is apparently known as "ringing". The colluding bidders can, for example, keep the price low and then share the profit. It is a strong statement that this is always possible but the statement is relative to the protocol level. In the formal modelling and analysis of the auction protocol using the inductive approach we will see that they are right. The challenge is to extend the usual protocol model with context information so that it becomes feasible to express this kind of collusion and consequently formalize and proof stronger security properties.

The second attack on the cocaine auction protocol is the so-called "sweetheart deal": the collusion between the seller and one of the bidders, i.e., "seller not selling to the highest bidder" in [25, p. 4]. If this attack is attempted within the limits of the protocol, i.e., the seller sends the secret appointment not to w_{n-1}, i.e., the winner of the winning round $n-1$, but instead to one of the earlier bidders w_i for $i < n-1$, then the protocol in the above implementation with Diffie-Hellman fails. That is, if S sends message n encrypted with key $g^{x_i y} \bmod p$ instead of $g^{x_{n-1} y}$, the real winner w_i will not be able to decrypt the message. This failure of the protocol means that the attack is infeasible if the protocol is implemented correctly. Formalizing the protocol should enable proving that this is the case. We will see how this can be formally proved even on our own implementation with public keys in Section 4.

The attacks on the cocaine auction protocol involve bidders and sellers. They are attacks on the auction that are only possible because they exploit privileges, like knowledge of keys, certificates, and access rights of roles, granted to peers in the protocol. Therefore, they can be categorized as insider attacks.

4. FORMAL MODEL

The Diffie-Hellman key exchange is a very efficient implementation of this protocol. However, we aim at using the established inductive approach to formally verify the protocol. Unfortunately, the inductive approach uses an abstract specification of symmetric or asymmetric keys and the Diffie-Hellman keys do not fall in those categories. In fact, the established Diffie-Hellman key $g^{x_i y} \bmod p$ is a symmetric key while the so-called ephemeral keys $g^y \bmod p$ and $g^{x_i} \bmod p$ are no encryption keys rather intermediate computations encrypting the secrets y and x_i for public exchange. Therefore, we provide here another possible implementation of the cocaine auction protocol using standard public-key cryptography.

The cocaine auction protocol can then be specified using the public key K_S of the seller S and its secret counterpart K_S^{-1} for decryption and public encryption keys K_{A_i} of the bidders with corresponding secret decryption keys $K_{A_i}^{-1}$.

0. $\mathrm{S} \to \mathcal{D}$: K_S

i. $?A_i \to \mathcal{D}$: $\{K_{A_i}, b_i\}_{K_S}$

n. $\mathrm{S} \to \mathcal{D}$: $\{b_{n-1},\ \mathit{MeetingAppointment}\}_{K_{A_{n-1}}}$

This asymmetric version of the protocol works as follows.

- In Step 0, the seller sends to all bidders in the set $\mathcal{D}$ a public key K_S enabling them to send secret message to the seller.
- In each of the rounds i for $i = 1, \ldots, n-1$ the bidder sends anonymously using the sender address $?A_i$ his public key for the round K_{A_i} together with the prearranged bid b_i for the round encrypted with the public key of the seller K_S to the seller. The contents of this message are only visible to the seller since only he holds K_S^{-1}.
- In the final round (after the timeout has happened) the seller S broadcasts to all bidders in $\mathcal{D}$ the highest bid b_{n-1} and the secret message with the *MeetingAppointment*. This broadcast message is encrypted with the public key $K_{A_{n-1}}$ of the winner of round $n-1$ that the seller could retrieve from the message in the winning round $n-1$.

In comparison to the version using Diffie-Hellman key-exchange, this second implementation of the cocaine auction protocol seems slightly more complex. Although the latter implementation abstracts from a concrete public-key algorithm, Step i requires an encryption of the entire message $\{K_{A_i}, b_i\}$, whereas the Diffie-Hellman version requires only the computation of one ephemeral key $g^{x_i} \bmod p$. For example, if we consider RSA to be the concrete algorithm used in the second public key version, the key length, i.e., the size of the modules p would be roughly the same for RSA and Diffie-Hellman for equal strength of security. At a closer look, however, computing $g^{x_i} \bmod p$ corresponds roughly to the same computation effort as computing $(\#\{K_{A_i}, b_i\})^{K_{A_i}} \bmod n$ (following the RSA-algorithm [24] where the # is the transformation of the message into a number for exponentiation with the RSA encryption key K_{A_i} modulo the public modulus $n = p * q$). The two implementations, or more precisely Steps i are of similar complexity, because, the ps and qs are of similar size (currently 1024 bits are still considered safe, although 2048 are recommended after the successful factorization of a 1024 prime equivalent to breaking of 700 bit RSA key and the Logjam attack on Diffie-Hellman [2]).

4.1 Isabelle's Inductive Approach to Security Protocol Verification

The interactive theorem prover Isabelle/HOL [21] implements classical higher order logic (HOL) for the modelling of application logics. Inductive definitions and datatype definitions can be written in a way close to programming languages. Semantic properties over datatypes can be formalized in a simple equation style by primitive recursion and are strongly supported by automated proof procedures based on rewriting, automated simplification, as well as externally coupled dedicated provers.

The inductive approach to security protocol verification by Paulson [23], the designer of the Isabelle system, picked up on the hype generated by the earlier model checking approach to security by Lowe [16]. In comparison, the inductive approach is more laborious as it requires human interaction, but it is unrivalled in its expressiveness which allows proofs beyond the ones that are usually done in model checkers. Although proofs in Isabelle/HOL are not performed automatically but have to be provided by the user, the increased expressiveness allows modelling protocols less abstractly than in a model checker. A protocol's definition is

given as an inductive definition of the set of all "traces" that are allowed by the protocol. A trace is a list of events representing the sending and receiving of messages that happen in a possible run of the protocol. The inductive definition defines all possible behaviours of a protocol as the minimal set of traces described by the inductive rules corresponding to the protocol's communication steps.

Paulson's inductive approach is only a starting point for the modelling of insider threats to auction protocols. We see in this paper, that the classical inductive approach needs to be extended with concepts developed for the Isabelle insider framework [12] in order to fully support reasoning about insiders. Since an Isabelle framework, like the inductive approach to Security Protocol Verification, is nothing other than a number of theory files containing a set of tailor-made definitions and related theorems, it can be easily extended and also integrated with other approaches like the Isabelle insider framework [12].

For the sake of self-sufficiency, we briefly present the main features of Isabelle's inductive approach concentrating on the parts we use.

4.1.1 Cryptography, Keys, and Messages

Security protocol specifications are constituted as sequences of communication steps between principals possibly adding abstract cryptographic functionality. For example, in the so-called "Alice-Bob"-notation a typical protocol step like

$$A \mapsto B : \{M\}_K$$

would read as: "A sends to B message M encrypted by key K". The key could be specified more precisely as a symmetric key or the private K_A^{-1} or public K_A key of an agent A.

The function `invKey` maps a public key to its matching private key, and vice versa.

```
type synonym key = nat
consts invKey :: key ⇒ key
```

Nonces are a means to avoid replay attacks. A nonce is a large random number. A message that requires a reply can incorporate a nonce. The reply then must include that same nonce to prove that it is not a replay of a past message.

Protocol messages can then be defined as a recursive datatype `msg` building over the simple message constituents agents, numbers, nonces, and keys. Protocol messages usually consist of more than just one component. For the recursive cases, a `msg` can be a combination of other messages or a message encrypted with a key.

```
datatype msg = Agent agent
             | Number nat
             | Nonce nat
             | Key   key
             | MPair msg msg
             | Crypt key msg
```

Isabelle offers a sophisticated pretty printing syntax facility. This allows us to define the notation `{`$x_1, \ldots, x_{n-1}, x_n$`}` for the nested pairing `MPair` $x_1 \ldots$ (`MPair` x_{n-1} x_n) making the specifications of protocols very close to the on-paper notation.

The way datatypes are implemented in Isabelle provides the property that all of datatype constructors are injective functions. Therefore, the above definition `msg` implicitly entails the following theorem.

```
Crypt K M = Crypt K' M' ⟹ K = K' ∧ M = M'
```

This theorem says that a message `M` encoded with a key `K` yields only one ciphertext `Crypt K M` and no other message `M'` can be mapped onto this ciphertext – not even with a different key. The model is an oversimplification: in reality decryption with a wrong key `K'` would actually yield a result although quite likely pure rubbish. The oversimplification is justified as in reality checksums are introduced on the plaintext to exclude decryption with wrong keys.

4.1.2 Attacker Model, Events, and Traces

The principals are expressed by a datatype definition guaranteeing their distinctiveness. We assume a server, a number of friendly principals, and a spy. That is, in our model the attacker is explicitly modelled.

```
datatype agent = Server | Friend nat | Spy
```

The attacker can forge messages using all components he can derive from previous traffic. The inductive operators characterize the constituents of a protocol's messages (set `parts`), messages the attacker can extract from a protocol trace (set `analz`), and messages that the attacker can build (set `synth`).

Protocols are defined by inductive definitions describing the behaviour of principals taking part in the protocol. Behaviours are sets of possible event traces. A trace is a list of communication events, such as interleaved protocol runs.

Compared to the inductive definitions for `synth` and `analz`, protocol definitions are thus of a different type: rather than specifying a message set, they specify the behaviour of the communicating principals as traces of *events* defined as a datatype comprising different cases (represented as datatype constructors) of protocol communication events. The main case of an event is that an agent sends a message to another agent: the constructor `Says` takes three arguments of types `agent`, `agent`, and `msg` and returns one result of type `event`. The other constructors of the datatype `event` are `Gets` and `Notes` to specify the reception and storing of messages. Defining a protocol in the inductive definition consists of defining a set of traces of events representing all possible runs of the specified protocol. The attacker's behaviour is added by including `Fake` messages into traces. The analysis first derives the knowledge he can extract from the protocol (`analz`) and the messages he can synthesize (`synth`). This characterizes the attacker's behaviour and allows verification of security properties. Following the Dolev-Yao model [7], the `Spy` gets to know everything that is communicated along any channel. To this end, the inductive approach models a function `spies` that effectively deconstructs event traces into sets of messages.

We omit this definition because we concentrate on insider threats, i.e., we cannot assume to have a clear distinction into "good" and "bad". Our attacker could be any agent. Besides insider threats, the cocaine auction protocol reveals other requirements to the inductive approach that go beyond classical security protocols.

1. In general, auctions necessitate an *arbitrary number of rounds*;
2. we need to represent *broadcast communication*;
3. we need to enable *anonymous sending* of messages.

4.2 Cocaine Protocol in Isabelle

Our formalization of the cocaine auction protocol resides in the theory file `CocaineAuction.thy` which is available online [8]. We provide next the inductive definition before we illustrate it by a simple example trace.

Formally, the inductive definition starts by introducing a constant `cocaine_auction`.

```
inductive_set cocaine_auction :: event list set
```

Following this introduction of the inductive set constant, a series of rules determines exactly which traces, i.e., lists of events, are in the set defining the semantics of the protocol. First, the rule `Nil` describes that the empty set is a possible trace, representing the beginning of each protocol run.

```
Nil: [] ∈ cocaine_auction
```

Similar to the specification of other protocols, we specify that `Spy` (see its specification in `Message.thy` [8]) can analyze and synthesize from what he "spies", i.e., the set of things he knows (see also `Event.thy` [8]). Spy can then say all these things since he is an agent as well. The symbol $\Longrightarrow$ is the right associative implication of Isabelle's meta logic, i.e., the first two sub-formulas below have to be read as a conjunction. The symbol # is the list constructor. Altogether, the following rule reads as "if a trace `evsf` is a (possible behaviour of a) cocaine auction *and* `X` is a message that can be synthesized from what can be analyzed from all the events in the trace that the spy can see, *then* a possible continuation is the sequence `evsf` extended by the event in which the spy utters to any principal `RR` this message `X`".

```
Fake:  evsf ∈ cocaine_auction
       ⟹ X ∈ synth (analz (spies evsf))
       ⟹ Says Spy RR X # evsf ∈ cocaine_auction
```

Initially at the beginning of every cocaine auction, the `Server` (equal seller) sends his public key out to all agents that are present (all `Friends`). This definition uses list comprehension to produce a list of `Says` events for all `i < friends` setting it as the beginning of any run of the cocaine auction.

```
CA0: [Says Server (Friend i) (Key(pubK Server)).
                  i ← [0..<friends]] ∈ cocaine_auction
```

In each round i, a `Friend` (bidder) can make an offer by saying "yes" corresponding to broadcasting a public encryption key encrypted with the Server key. Together with this public encryption key the bidder sends also the current price of that round (`bid i`) assumed to be given in advance by the function bid applied here to the round number `i`. We use the Isabelle `specification` device that allows to define an abstract function `bid` specifying that it should be injective, strongly monotonically increasing, and `bid(0) = 0`. A witness has to be given and the properties must be proved for a specification to be accepted by Isabelle. Authentication of the bidder is omitted here since later the bidder authenticates himself at the meeting point which he would not have been able to find without decrypting the message of the Server (see below – the final message of the Server is encrypted with the public key of the bidder transmitted here)). The bidder uses the sender address `Friend(friends)` which is the "anonymous" address. I.e., the bidders use "spoofing" to anonymize their messages. The public key they send is here formalized as `pubK(Friend j)` (see the theory file `Public.thy` [8]) but the owner of the key `Friend j` is assumed to be not visible – not even to the intended receiver of this broadcast message (the Server). The precondition on `hd(evs)` ensures that either

- this is the first round indicated by the last message (first in event list) being one of the initial messages of the Server with his `pubK`, or
- the previous event has been a message in which a bidder different from `Friend j` has made a bid and won the round. This is indicated in the last message being a bid similar to the current one but from `Friend k` with the previous `bid(i - 1)`. `Friend j` now can increase the price to `bid i`.

```
CAi:  evs ∈ cocaine_auction ⟹ j <  friends ⟹
  hd(evs) = Says Server (Friend l)(Key(pubK Server))
              ∧ i = 1  ∨
  hd(evs) = Says (Friend friends) Server
              (Crypt(pubK Server)
                {Key(pubK(Friend k)), Number(bid (i-1))})
              ∧ i > 1
  ⟹ (Says (Friend friends) Server
            (Crypt (pubK Server)
              {Key(pubK(Friend j)), Number(bid i)}))
       # [Says (Friend friends) (Friend k)
            (Crypt (pubK Server)
              {Key(pubK(Friend j)), Number(bid i)})).
                  i ← [0..<friends]]
       @ evs ∈ cocaine_auction
```

Timeout can take place at any time. We simply do not model time in the protocol and it is not necessary. Since we keep all possible traces as the semantics of a protocol in the inductive approach, any occurrence of timeouts is modelled. For the next rule of the auction protocol, we assume that timeout has just happened. Now, in this final round, the Server sends out the message with the secret appointment (encoded as a natural number for simplicity) and signs it with the public key of "some friend". This is the bidder that has won the previous round `n-1`. In this refined version, we enforce this by the precondition on `hd(evs)`. The winner of the previous round is represented in this most recent message by its public encryption key `pubK (Friend j)`. For the final message, this key of `Friend j` is chosen and the message with the secret appointment `mtng`[3] is sent encrypted with this public encryption key `pubK` of `Friend j` so that only he can decrypt it.

```
CAn: evs ∈ cocaine_auction ⟹
  evs = Says (Friend friends) Server
          (Crypt (pubK Server)
            {Key(pubK(Friend j)), Number(bid i)})
           # [Says (Friend friends)(Friend k)
                {Key(pubK(Friend j)), Number(bid i)}.
                   k ← [0..<friends]] @ evs
  ⟹ [Says Server (Friend k)
          (Crypt (pubK(Friend j))
            {Number(bid i), Number mtng}.
              k ← [0..<friends]]
       @ evs ∈ cocaine_auction
```

This specification of the cocaine auction protocol establishes a few particular solutions extending the inductive approach to represent the peculiar requirements of the application (see previous section).

[3]The message is for simplicity embedded into the given message type `msg` provided in the inductive approach. It may be thought of as "encoded" as a number.

1. *Arbitrary numbers of rounds* in an auction are enabled and yet inconsistent traces are excluded since the rule `CAi` can be chained up any number of times but using a natural number counter `i` their interleaving can be controlled. The use of `i` and `i-1` is based on the mathematical library of Isabelle showing yet again the advantage of using this expressive, complete and consistent approach.

2. *Broadcast communication* is modelled explicitly using lists of messages to all principals. Again we see here the use of Isabelle libraries – this time for lists using the list comprehension in Haskell-like syntax: the formalization of a broadcast of a message `m` from a principal A to a community of principals $\mathcal{D}$ is `[Says A (Friend j) m. j ←[0 ..<friends]`. This is simple and concise and corresponds quite closely to the specification using Alice-Bob notation as specified in the original paper [25] (also see Section 3.3).

3. *Anonymous sending* is implemented in our above protocol specification by *spoofing*. This term corresponds to a classical vulnerability of the TCP/IP protocol whereby the sender field in IP-packets can be freely replaced by an attacker to impersonate a principal. In order to hide his real identity, here in our inductive definition, the legitimate sender inserts (*spoofs*) the sender `Friend friends` in rule `CAi` using an identity that is out of bounds (only addresses strictly less than `friends` are admitted for `Friend`s).

Although we stated above that we "abstract from a concrete implementation of the anonymity layer" the implementation by spoofing discussed in Point 3 comes very close to a technical solution of an anonymity layer. However, it does implicitly use the context assumptions of Paulson's inductive approach, here specifically, that keys remain unbroken, and that no attacker has a complete view of the network, but also others that the attacker can intercept, eavesdrop, and insert fabricated messages. These assumptions are common as global assumption for security protocol verification. They are mainly due to the Dolev-Yao model but are also inspired by common properties of the Internet protocol TCP/IP, like the spoofing property used. Clearly, in the context of insider threats we would need a slightly more global view as we consider not only the networking layer but also higher layers of infrastructures, like physical architectures, organizational policies, and even socio-technical system aspects [12].

4.2.1 Simple Example Trace

In order to illustrate the inductive definition of the cocaine auction protocol we consider here a simple example. Assume there are only 2 bidders, i.e., `friends = 2`. The following subset of traces of `cocaine_auction` step-by-step grows traces representing an auction in which each bidder makes just one offer before timeout appears after `Friend 1` bids finishing the auction. In the following set lists are postfixes of their successors, i.e., the traces repeat the previous trace. To make the exposition more succinct, we put "..." as much as possible omitting repetitions but their last element and highlighting the common parts by equal colours. There are precisely two rounds. We omit in particular all traces interleaved by `Fake` events.

```
{
[],
[ Says Server (Friend 0)(Key(pubK Server)),
  Says Server (Friend 1)(Key(pubK Server))
],
[ Says (Friend friends) Server
       (Crypt(pubK Server)
          {Key(pubK(Friend 0)), Number(bid 1)}),
  Says (Friend friends) (Friend 0)
       (Crypt(pubK Server)
          {Key(pubK(Friend 0)), Number(bid 1)}),
  Says (Friend friends) (Friend 1)
       (Crypt(pubK Server)
          {Key(pubK(Friend 0)), Number(bid 1)}),
Says Server (Friend 0)(Key(pubK Server)),
  Says Server (Friend 1)(Key(pubK Server))
],
[ Says (Friend friends) Server
       (Crypt (pubK Server)
          {Key(pubK(Friend 1)), Number(bid 2)}),
  Says (Friend friends) (Friend 0)
       (Crypt(pubK Server)
          {Key(pubK(Friend 1)), Number(bid 2)}),
  Says (Friend friends) (Friend 1)
       (Crypt(pubK Server)
          {Key(pubK(Friend 1)), Number(bid 2)}),
Says (Friend friends) Server
       (Crypt(pubK Server)
          {Key(pubK(Friend 0)), Number(bid 1)}),
  ...
],
[ Says Server (Friend 0)
       (Crypt(pubK(Friend 1))
          {(Number(bid 2)), Number 42)},
  Says Server (Friend 1)
       (Crypt(pubK(Friend 1))
          {(Number(bid 2)), Number 42)},
  Says (Friend friends) Server
       (Crypt (pubK Server)
          {Key(pubK(Friend 1)), Number(bid 2)}),
  ...
]
}
```

Clearly this is just an illustrative small example given by a selected subset of traces following the provided inductive rules in the given order `CA0, CAi, CAn`. The message that the Server sends in the final step encoding the meeting appointment as a number is randomly chosen to be 42; uniquely encoding a real appointment message like "meet me at 6.30am in the car park of Heathrow Terminal 5" would result in a much larger number.

4.3 Specification and Proof of Insider Threats

The insider attacks on auctions we investigate on the running example of the cocaine auction protocol are the sweetheart deal and the collusion of bidders known as "ringing" because they build a bidding ring or cartel (see Section 2).

4.3.1 Sweetheart Deal

Our hypothesis is that the formal specification of a security protocol is sufficient to exclude the sweetheart deal. That is, the way we defined the rules for the cocaine auction should forbid that the seller announces the wrong bidder as the winner (his "sweetheart" – someone he has made a deal with outside the auction).

As a first observation, this attack is clearly an insider attack, as it it only possible because an insider – here the seller – colludes with another insider – a bidder. Together they use their privileges given by the policy – here, the auction –

to achieve the attack goal – here, winning the auction.

The second observation is that the two final steps of the protocol – the way we defined it – prohibit that this insider threat may occur. In our formal specification, the second to last step of any protocol run (not counting interspersed Spy actions) is an application of `CAi` (see also the example given in the previous section to illustrate that point). The second to last message is a broadcast message of the bidder `Friend j` to the Server and all other bidders using the anonymous sender address `Friend friends` but containing an own public key `pubK(Friend j)` encrypted with the Server's public key. A list of events corresponding to this broadcast message must be starting the trace if the last rule `CAn` is invoked. When applying the rule, the last broadcast messages thus automatically use the key `(Crypt (pubK(Friend j))` for the encryption of the meeting appointment for that same `Friend j` as specified in the precondition. Therefore, no other `Friend k` for `j` $\neq$ `k` can be chosen by the Server.

Informally, this argument seems clear. But how can we prove this formally? The first step is the statement of the property which just formalizes the above observation. If any cocaine auction ends with a broadcast by the Server that the bidder `Friend j` is the winner, then the trace `evs` prior to this must have been a broadcast of this bidder. The additional assumption `0 < friends` in the theorem just excludes the empty set of bidders and generalizes the property for any finite number of bidders.

```
theorem no_sweetheart_deal:
0 < friends ⟹
[Says Server (Friend k)
      (Crypt (pubK (Friend j))
         {Number(bid i), Number mtng}.
            k ← [0..<friends]]
@ evs ∈ cocaine_auction
⟹ ∃ evsf. evs =
  Says (Friend friends) Server
    (Crypt (pubK Server)
       {Key(pubK(Friend j)), Number(bid i)})
  # [Says Server (Friend k)
       (Crypt (pubK Server)
          {Key(pubK(Friend j)), Number(bid i)}).
              k ← [0..<friends]]
  @ evsf
```

Isabelle is an interactive theorem prover, i.e., statements of theorems, like the above, need to be proved. This proof is supported by the fact that the rules for defining the protocol are an inductive definition. In Isabelle, and also in general, inductive definitions define the least set that is closed by a given set of rules. The principle of *rule inversion* allows us for a given element in this set to make a case analysis according to the cases defined by the rules of the inductive definition. In our case, the elements of the inductive set are traces, i.e., lists of events. Applying rule inversion technically in Isabelle is provided by the command `inductive cases` which provides us with a case analysis rule that reduces a property statement about trace sets of cocaine auctions, like the theorem `no_sweetheart_deal` to 6 subgoals corresponding to the premises of each of the rules of the inductive definition.

For the proof of the theorem, luckily, the first `empty` case is trivially true, while 4 other cases can be easily excluded. In Isabelle, elements of a datatype, here the datatypes of events, message, and agents, are distinct if their arguments or constructors differ. Thus, two traces starting say with `Says Server X y` and `Says (Friend j) Z U` can never be equal because the arguments `Server` and `Friend j` are distinct therefore the application of constructor `Says` renders distinct elements. Consequently, the only real case that remains to be shown is the one that actually corresponds to the precondition of the theorem.

```
∃ evsf. evs =
  Says (Friend friends) Server
      (Crypt (pubK Server)
         {Key (pubK(Friend j)), Number(bid i)})
  # [Says (Friend friends)(Friend k)
       (Crypt (pubK Server)
          {Key(pubK(Friend j),Number(bid i)})).
             k ← [0..<friends]]
  @ evsf
```

This case can be easily solved by instantiating the existential quantifier and applying simplification.

4.3.2 Intermediate Analysis

An important observation from the previous attack is that the main attack analysis device of the inductive approach – the Dolev-Yao attacker `Spy` and the related infrastructure – play almost no role in it: possible injections of `Spy`-events into successful, i.e., finishing, traces of the cocaine protocol are merely those where the Spy feeds messages after Step 0 of the cocaine protocol. Even though `Spy` is able to play in `Fake` messages at any point of a partially finished trace this will lead to this trace ending unsuccessfully without reaching the goal of the auction. In the formal model, this is due to the fact that all latter steps require as a precondition that the previous message was one either originating from the `Server` or one from one of the `Friends j` for `j < friends`. Now, since `Spy, Server`, and `Friend j` are elements of the datatype `agent` that are created by different constructors, they are pairwise distinct. In particular, `Spy` cannot match either `Server` or `Friend j` for any `j` and the preconditions for any of the rules, `CAi` or `CAn` cannot become true any more once `Spy` has interspersed a trace by sending a fake message.

This limitation of the inductive approach is not surprising since modelling the agents as constructors of a datatype feeds into the global view (already discussed above) that agents are firmly divided into "bad" and "good".

Surprisingly, this does not impede the analysis of the sweetheart deal. On the contrary, for an analysis of insider threats in general, the fixed distinction of attackers and "good" principals is generally an inadequate modelling decision. As already observed in earlier papers on the formal analysis of insider threats [11], one of the major tricks to find attacks on security protocols is to consider insiders: the classic attack on the Needham-Schroeder attack is performed by the insider Eve[4]. This man-in-the-middle attack (or mirror-attack) uses impersonation which has motivated the Isabelle insider approach of using sociological model inspired by Max Weber supported with Hempel and Oppenheim's logic of explanation to model and analyze insider threats in Isabelle with logic and proof.

Therefore, at this point we extend the inductive approach with the Isabelle insider framework that has been especially

[4]Eve uses her own legal credentials (a public key) to get Alice's nonce sent to Eve when Alice wants to communicate with her. This nonce is used as an authentication token to Bob. Eve next uses the first protocol run with Alice as an oracle to decrypt Bob's Nonce sent back to Eve encrypted with Alice's public key to challenge her presumed identity.

designed for the purpose. We only introduce the minimally necessary parts of that framework in order to illustrate how it can be used to model ringing. For more detail, the interested reader is referred to the main paper [12] and application examples to IoT insider threats [10] and insider threats to Airplane safety and security [9]. The following section recapitulates the parts of the Isabelle insider framework that are used to extend the inductive approach because they are needed to express collusion (see theorem `Insider_homo_oeconomicus` below).

4.3.3 Isabelle Insider Framework

The Isabelle insider framework uses a taxonomy of insider threats [22]. This taxonomy is based on a thorough survey on results from counterproductive workplace behaviour, e.g., [19, 17] and case studies from the CMU-CERT Insider Threat Guide [4]. The insider framework simply models the taxonomy in HOL as datatypes, a concept of HOL that resembles the concept of taxonomy classes. As an example, consider the formal representation of *Psychological State* [22] as a datatype.

```
datatype psy_states = happy | depressed | disgruntled
                    | angry | stressed
```

The element on the right hand side are the five injective constructors of the new datatype `psy_states`. They are simple constants, modelled as functions without arguments. Another example is *Motivation* [22].

```
datatype motivations = financial | political | revenge
                     | fun | competitive_advantage
                     | power | peer_recognition
```

In the Isabelle insider framework, we combine the characteristics about the actor in a combined state.

```
datatype actor_state = State motivation psy_state
```

The *Precipitating Event* or *Catalyst* can be any event that has the potential to tip the insider over the edge into becoming a threat to their employer. It has been called the 'tipping point' in the literature. This catalyst is encoded as a tipping point predicate describing the psychological state and motivation of an actor to become an insider.

```
definition tipping_point :: actor_state ⇒ bool
tipping_point a ≡ motivation a ≠ {}
                ∧ happy ≠ psy_states a
```

Insider threat case studies show that a recurring scheme in insider attacks lies in role identification as described in [11]. The Isabelle insider framework uses this role identification in the definition of the `UasI` predicate. It expresses that the insider plays a loyal member of an organization while he simultaneously acts as an attacker. Note, that in order to integrate the Isabelle insider framework with the inductive approach, we use the `agent` constructor `Friend` here.

```
UasI a b ≡ (Friend a = Friend b)
```

Insider attacks link the insider characterization of psychological disposition with the above insider behaviour `UasI`. This is defined by the following rule `Insider a C` for the attacker `a`. The parameter `C` is a set of identities representing the members of an organization that are to be considered as safe.

```
Insider a C ≡
tipping_point (astate a) ⟶ (∀ b ∈ C. UasI a b)
```

Although the above insider predicate is a rule, it is not axiomatized. It is just an Isabelle definition i.e., it serves as an abbreviation. To use it in an application, like the auction protocol, we can use this rule as a local assumption in theorems (see below theorem `Insider_homo_oeconomicus`) or using the `assumes` feature of locales [13]).

4.3.4 Homo Oeconomicus and Ringing Attack

The principle of *homo oeconomicus* defines agents to be rational. In general, this economic principle captures the idea that any agent a will not spend more than necessary to get an asset, i.e., if a can get the asset for price X, he will not pay price $Y > X$ to get it.

Without explicitly introducing the additional concept of "price" and buying assets, we can simply formalize the principle `homo_oeconomicus` for the context of the cocaine protocol, by stating that an agent that is currently the winner of round `i` will not make another bid in the next round. Technically, as a general property this states that for all traces `t` representing (intermediate) runs of the cocaine protocol no bidder will make a bid in the current round, if he is the highest bidder in the trace leading up to the current round. This is represented by the function `CAtl` applied to cocaine auction trace `t`. We define this function `CAtl` as a primitive recursive function that cuts off from any trace `t` all leading events if these exist at the front of `t` corresponding to (a) the Server's final broadcast according to rule `CAn`, (b) the last bid, i.e., the anonymous broadcast by some `Friend j` according to rule `CAi`, (c) all initial Server messages according to rule `CA0` leaving the empty trace. Excluding that the currently highest bidder will make the next bid, corresponds to saying that the *head* of any trail `t` cannot be an event in which this bidder broadcasts his "yes". We can simply use the Isabelle list function `hd` since literally the first element of that list of "yes" broadcast events corresponding to rule `CAi` is the message to the Server. The auxiliary functions `highest_bidder` and `cur_round` are also defined as primitive recursive functions over traces in Isabelle in the intuitive way (for details see the Isabelle files [8]).

```
homo_oeconomicus ≡
∀ t ∈ cocaine_auction. ∀ j < friends.
  highest_bidder (CAtl t) (Friend j) ⟶
    hd t ≠ Says (Friend friends) Server
      (Crypt (pubK Server)
         {Key(pubK(Friend j)), Number(bid(cur_round t))})
```

As a first illustration for the use of this definition, we can use it to show that if there is only one bidder, the seller will only get the reserve price `bid 1`.

```
theorem homo_economicus_one_bidder:
friends = 1 ⟹ homo_oeconomicus
⟹
∀ t ∈ cocaine_auction.
t = [Says Server (Friend k)
      (Crypt(pubK(Friend j)){Number(bid i), Number msg}).
        k ← [0..<friends]]
    @ evs
⟶ i = 1
```

The above property is a useful stepping stone on the way to proving that if there is a collusion amongst all bidders, the Server will only get the reserve price. We cannot prove that a collusion between players glues them together to become physically one `Friend` corresponding to showing that the constant `friends` must be equal to 1. However, we can

prove that the same conclusion as in the previous theorem follows as well. We assume that one bidder, `Friend 0` is an insider and at the tipping point, and all bidders *act* as one agent, i.e., the insider can impersonate them. From that we show the same conclusion as in the previous theorem follows: the seller only gets the reserve price.

```
theorem Insider_homo_oeconomicus:
homo_oeconomicus ⟹ tipping_point(astate 0)
⟹ Insider 0 {i. i < friends} ⟹
∀ t ∈ cocaine_auction.
t = [Says Server (Friend k)
       (Crypt (pubK(Friend j))
          {Number(bid i), Number msg}).
            k ← [0..<friends]]
    @ evs
⟶ i = 1
```

In the proof of this theorem, the insider assumption is used to show that the prerequisite `highest_bidder (CAtl t) (Friend j)` of the hypothesis `homo_oeconomicus` can be made true for any bidder as soon as one of them, here `Friend 0`, is the highest bidder, because he can impersonate anyone. Invoking the assumption `homo_oeconomicus`, we can then prove that any continuation of the trace containing a first bid cannot continue, since it would be a bid of the same bidder contradicting the principle. So all traces end with the highest bid `bid 1` which corresponds intuitively to a "reserve price" (`bid 0` is specified to be `0`, see Section 4.2).

We have thus formally shown that the insider assumption in fact enforces that the cocaine protocol can generally be corrupted by the collusion of all bidders.

5. USEFULNESS AND LIMITATIONS

In the protocol we make several assumptions. We share the view of the authors of the original paper [25] "We do not believe that all such attacks can be detected, let alone stopped, by any particular auction protocol".

One problem is inherent in the set-up of the auction. In order to avoid that the other participants see who has made a bid at a particular time, the authors of [25] discuss that the participants have a clicker in their pockets and can press them without the others noticing. Then there are some problems with this, since once a bid has been made another press of a button would mean to bid for the next higher price. For example, assume that in each step the price goes up by 1000 and that the current high bid is at 49,000. Assume furthermore that both bidder b_1 and b_2 are willing to bid up to (inclusively) 50,000. They both decide to press, but b_1 is a split second faster and makes the bid for 50,000. The click by b_2 is still registered but would count as 51,000, an amount b_2 would not want to pay. Practically it would make sense that after a click any further clicks are disabled by a fixed amount of time (e.g., 10 seconds) and the amount of the current high bid is announced (e.g., on a display). In this scenario b_1's bid would initially disable the bidding process and after the display of 50,000 b_2 would have to press again before a further bid is registered. We assume that this process cannot be manipulated, since otherwise the auctioneer could always broadcast the acknowledgement of the bid with his sweetheart's key and all other bidders assume that they had been too slow to win the round, although actually one of them should have won the round and the sweetheart did actually not bid. The participants could not detect the manipulations since the actually generated trace is a legal trace of the protocol.

The proof guarantees only that the key used by the winner of the penultimate round – after the 30 seconds have lapsed without a bid – is used for the broadcast with the secret location, so that only the winner can decrypt the location. If, however, the true winner did not know that he had won the penultimate round, he would not expect to be sent the location and would not be in the position to detect foul play.

It is not surprising that without any assumptions only very little can be said about possible traces. In this case, it can be said that it can be detected if different keys are used in the penultimate and in the ultimate rounds. Obviously, the protocol cannot rule out that the seller sends a wrong *MeetingAppointment* to the true winner and uses communication channels outside the protocol to communicate the true *MeetingAppointment* to his sweetheart.

In summary, our model abstracts from some implementation details and inherently assumes the following.

- The implementation of the auction needs to provide a mechanism to avoid racing conditions and give unambiguous feedback to the successful bidder, e.g. some notice board and time delays between bids.
- The veracity of the meeting point is assumed and the post-procedure of the cocaine-money exchange is beyond the protocol model.

In fact, the part of the Isabelle insider framework that has not been used here provides the possibility to express infrastructure in which agents act and could thus be used to address the second point. However, that would be beyond the limits of this paper.

An interesting question is, how in the inductive approach, the attack on the Needham-Schroeder asymmetric protocol (NS-protocol) has been modelled. This man-in-the-middle attack can be considered as the first insider attack [11]. However, in the inductive approach the attacker is always only the agent `Spy`, and `Spy` is different to all other agents by construction. So, how could the NS-protocol be modelled when the attacker is `Friend i` for some `i` in `n`? The answer is that the protocol definition deliberately allows any agent in the rules for the inductive definition. In the rules of the definition of the NS-protocol in the inductive approach, letters `A` and `B` are used to suggest agents Alice and Bob. Consider, for example the crucial rule `NS1` from the NS-protocol formalization in the inductive approach where the initiator sends a nonce to the intended recipient.

```
NS1:  evs1 ∈ ns_public ⟹  Nonce NA ∉ used evs1 ⟹
       Says A B (Crypt (pubEK B) { Nonce NA, Agent A})
           # evs1  ∈  ns_public
```

The important point for the NS-protocol attack to work is that this intended recipient can be an attacker. I.e., a legal participant of the network of peers is malicious and abuses a connection request by the initiator to impersonate this initiator. Although in the above rule, the capital letters `A` and `B` seem to indicate that these are the agents Alice and Bob, they are variables fixed only in the context of the rule. When the rule is applied they can be freely instantiated to any agent, also to `Spy`.

Summarizing, the inductive approach allows for different "kinds" of specification of a protocol: one where the actors are explicitly made distinct using different constructors of datatype agent in the specification (this is how we used it

for the cocaine auction protocol) and one where agents are all abstract within the protocol (either as Higher Order variables as in the NS-application above or all represented as `Friend j`.) If we were to redefine the cocaine auction protocol in the latter way with abstract actors for all roles, then we would replace the seller also by `Friend k`. In this case, we would not be able to exclude the sweetheart deal: if we assume that the Server, say `Friend 0`, is an insider and at tipping point, he can impersonate a bidder and can then make his own bids using another role, say Friend j. This would enable the Server to provide a suitable bid for his sweetheart. In addition, it would enable another attack in which the Server, acting like a bidder, just drives the price up.

The formalization and proofs presented in this paper provide in summary the following results.

- Formal model of the cocaine auction protocol using the inductive approach proving the absence of sweetheart deals and the impossibility to exclude collusion.
- Formalization of arbitrary numbers of rounds, broadcast, and anonymous message sending for the inductive approach.
- The inductive approach can only deal with Insider threats by abstracting from its agent datatype that prevents good agents from behaving badly.
- Integrating (parts of) the Isabelle insider framework with the inductive approach enables reasoning about collusion of insiders for auctions.
- The collusion exhibits that the assumption *homo oeconomicus* suffices to prove that rational insiders may use collusion to force the reserve price.

6. CONCLUSIONS

In this paper, we have investigated the vulnerability of auctions to insider attacks in particular different forms of collusion. We used the cocaine auction protocol as a case study that takes mistrust to an extreme. Using modelling and analysis in Isabelle we experimented with two different approaches, the inductive approach to security protocols and the Isabelle insider framework. We were able to model the protocol in the inductive approach and show that its formal specification excludes a possible insider attack, the "sweetheart deal". Integrating the inductive approach with the Isabelle insider framework enabled showing that collusion between all bidders, so-called "ringing", cannot be excluded. In order to prove the latter theorem, we formalized a notion of "homo-oeconomicus" for the cocaine auction protocol.

While some of the issues around broadcasting have been addressed in more recent extensions of the inductive approach [18] for group protocols, we provide our own solutions tailored to the specific needs of insider threats.

Limitations of our model are discussed in the previous section showing that – despite the formality introduced and frameworks used – all guarantees depend on the abstraction we chose when modelling. Any formalization and proof of system properties depends always on the model we consider. This is also true for the system abstractions of protocols and auctions. Therefore, the implicit assumptions about real world participants are crucial. The use of the insider framework makes role impersonation explicit in models and therefore helps to understand in more detail how insider attacks work in auctions. The additional assumption *homo oeconomicus* could be a beneficial extension to enrich the Isabelle insider framework by a notion of a rational insider although its general assumption as part of an Insider definition is disputable.

Acknowledgment

Part of the research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 318003 (TRE$_S$PASS). This publication reflects only the authors' views and the Union is not liable for any use that may be made of the information contained herein.

7. REFERENCES

[1] *Proceedings of the third IEEE Workshop on Research in Insider Threats, WRIT'14*. IEEE, 2014.

[2] D. Adrian, K. Bhargavan, Z. Durumeric, P. Gaudry, M. Green, J. A. Halderman, N. Heninger, D. Springall, E. Thomé, L. Valenta, B. VanderSloot, E. Wustrow, S. Zanella-Béguelin, and P. Zimmermann. Imperfect forward secrecy: How Diffie-Hellman fails in practice. In *22nd ACM Conference on Computer and Communications Security*, oct 2015.

[3] M. B. Caminati, M. Kerber, C. Lange, and C. Rowat. Sound auction specification and implementation. In *16th ACM Conference on Economics and Computation, EC'15*. ACM, 2015.

[4] D. M. Cappelli, A. P. Moore, and R. F. Trzeciak. *The CERT Guide to Insider Threats: How to Prevent, Detect, and Respond to Information Technology Crimes (Theft, Sabotage, Fraud)*. SEI Series in Software Engineering. Addison-Wesley Professional, 1 edition, Feb. 2012.

[5] D. Chaum. The dining cryptographers problem: Unconditional sender and recipient untraceability. *Journal of Cryptology*, 1:65–75, 1988. Unofficial copy at http://www.scu.edu/SCU/Programs/HighTechLaw/courses/ccp/diningcr.html.

[6] W. Diffie and M. Hellman. New directions in cryptography. *IEEE Transactions in Information Theory*, IT-22(6):644–654, November 1976.

[7] D. Dolev and A. C. Yao. On the security of public key protocols. In *22nd Annual Symposium on Foundations of Computer Science, SFCS '81*. IEEE, 1981.

[8] F. Kammüller. Isabelle formalisation of cocaine auction protocol., 2016. Available from https://www.dropbox.com/sh/uyku2q2ofb69cwm/AACYcxAnuI75YIOyqdz4wcoua?dl=0.

[9] F. Kammüller and M. Kerber. Investigating airplane safety and security against insider threats using logical modeling. In *IEEE Security and Privacy Workshops, Workshop on Research in Insider Threats, WRIT'16*. IEEE, 2016.

[10] F. Kammüller, J. R. C. Nurse, and C. W. Probst. Attack tree analysis for insider threats on the iot using isabelle. In *Human Aspects of Information Security, Privacy, and Trust - Fourth International Conference, HAS 2015, Held as Part of HCI International 2016,*

Toronto, Lecture Notes in Computer Science. Springer, 2016. Invited paper.

[11] F. Kammüller and C. W. Probst. Combining generated data models with formal invalidation for insider threat analysis. In *IEEE Security and Privacy Workshops (SPW)* [1].

[12] F. Kammüller and C. W. Probst. Modeling and verification of insider threats using logical analysis. *IEEE Systems Journal, Special issue on Insider Threats to Information Security, Digital Espionage, and Counter Intelligence*, 2016. Accepted for publication.

[13] F. Kammüller, M. Wenzel, and L. C. Paulson. Locales - a sectioning concept for isabelle. In Y. Bertot, G. Dowek, A. Hirschowitz, C. Paulin, , and L. Thery, editors, *Theorem Proving in Higher Order Logics, 12th International Conference, TPHOLs'99*, volume 1690 of *LNCS*. Springer, 1999.

[14] P. Klemperer. *Auctions: Theory and Practice*. Toulouse Lectures in Economics. Princeton University Press, 2004.

[15] V. Krishna. *Auction Theory*. Academic Press, 2002.

[16] G. Lowe. Casper: A compiler for the analysis of security protocols. In *Computer Security Foundations Workshop (CSFW '97)*, 1997.

[17] B. Marcu and H. Schuler. Antecedents of counterproductive behaviour at work: a general perspective. *Journal of Applied Psychology*, 89(4):647, 2004.

[18] J. E. Martina and L. C. Paulson. Verifying multicast-based security protocols using the inductive method. *International Journal of Information Security*, 14(2):187–204, 2015.

[19] M. J. Martinko, M. J. Grundlach, and S. C. Douglas. Toward an integrative theory of counterproductive workplace behaviour. *International Journal of Selection and Assessment*, 10(1–2):36–50, 2002.

[20] U. Maurer and S. Wolf. The Diffie-Hellman protocol. *Designs, Codes and Cryptography*, 19(3):147–171, Jan. 2000.

[21] T. Nipkow, L. C. Paulson, and M. Wenzel. *Isabelle/HOL – A Proof Assistant for Higher-Order Logic*, volume 2283 of *LNCS*. Springer-Verlag, 2002.

[22] J. R. C. Nurse, O. Buckley, P. A. Legg, M. Goldsmith, S. Creese, G. R. T. Wright, and M. Whitty. Understanding Insider Threat: A Framework for Characterising Attacks. In *IEEE Security and Privacy Workshops (SPW)* [1].

[23] L. C. Paulson. The inductive approach to verifying cryptographic protocols. *Journal of Computer Security*, 6(1-2):85–128, 1998.

[24] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, 21(2):120–126, Feb. 1978.

[25] F. Stajano and R. Anderson. The cocaine auction protocol: On the power of anonymous broadcast. In A. Pfitzmann, editor, *Information Hiding, Third International Workshop, IH'99*, volume 1768 of *LNCS*, pages 434–447. Springer, 1999.